FEBS Letters 583 (2009) 2547–2551

journal homepage: www.FEBSLetters.org

Predictive bioinformatic identification of minor receptor group human rhinoviruses

Christoph Weber [a], Angela Pickl-Herk [a], Abdul Ghafoor Khan [a], Sascha Strauss [b], Oliviero Carugo [c,d], Dieter Blaas [a,*]

[a] Department of Medical Biochemistry, Medical University of Vienna, Max F. Perutz Laboratories, Dr. Bohr Gasse 9, A-1030 Vienna, Austria
[b] CIBIV, Univ. Vienna and University of Veterinary Medicine Vienna, Austria
[c] Department of General Chemistry, University of Pavia, Viale Taramelli 12, I-27100 Pavia, Italy
[d] Department of Structural and Computational Biology, Max F. Perutz Laboratories, Vienna Biocenter, Vienna University, Campus Vienna Biocenter 5, A-1030 Vienna, Austria

ARTICLE INFO

Article history:
Received 15 May 2009
Revised 9 July 2009
Accepted 13 July 2009
Available online 16 July 2009

Edited by Gianni Cesareni

Keywords:
In silico
Molecular modelling
Rhinovirus
LDL-receptor
Energy calculation
Affinity
Bioinformatics

ABSTRACT

Major group HRVs bind intercellular adhesion molecule 1 and minor group HRVs bind members of the low-density lipoprotein receptor (LDLR) family for cell entry. Whereas the former share common sequence motives in their viral capsid proteins (VPs), in the latter only a lysine residue within the binding epitope in VP1 is conserved; this lysine is also present in "K-type" major group HRVs that fail to use LDLR for infection. By using the available sequences three-dimensional models of VP1 of all HRVs were built and binding energies, with respect to module 3 of the very-low-density lipoprotein receptor, were calculated. Based on the predicted affinities K-type HRVs and minor group HRVs were correctly classified.

1. Introduction

Human rhinoviruses, two established (A and B) and one proposed species (C) within the genus Enterovirus circulate as more than 100 types in the human population causing common colds. For the best characterized species A and B the receptors for host cell access are known; 12 types, the minor group, bind low-density lipoprotein receptor (LDLR), very-LDLR (VLDLR), and LDLR-related protein (LRP) and 87 types, the major group, bind intercellular adhesion molecule 1 (ICAM-1). All minor group HRVs are species A, whereas major group HRVs belong either to species A or B [1]. The recently discovered species C is poorly characterized biochemically and its receptor(s) is not known [2]. Recently, genome sequences of all known HRV serotypes and of several field isolates have been determined [3].

The genomic single stranded (+) RNA genome is enclosed within an icosahedral shell of 30 nm diameter composed of 60 copies of each of the capsid proteins VP1, VP2, VP3, and VP4. The binding sites of the respective receptors have been determined via electron cryo-microscopy and X-ray crystallography of complexes between virus and soluble receptor fragments. ICAM-1 binds within the canyon, a cleft encircling the fivefold axis of icosahedral symmetry [4] and contacts either of two motives conserved within each of the species (except from minor group HRVs that lack these motives) [5]. Conversely, as shown for the minor group virus HRV2, human VLDL-receptors attach via several of their ligand-binding repeats to the BC, DE, and HI loops of VP1 that build a star-shaped dome at the fivefold axis close to the icosahedral vertex [6]. Within the receptor footprint the 12 minor group HRVs only exhibit a common lysine residue at the tip of the HI-loop but the remaining residues are highly variable as they also contribute to the type-specific antigenic epitopes. Even when taking into account spatial vicinity within the three-dimensional structure, no obviously conserved amino acid pattern is apparent. An additional complication in understanding receptor recognition is the existence of 10 major group HRVs that also possess a lysine residue (and were therefore termed 'K-type HRVs') at a position equivalent to that of the lysine in minor group HRVs. Based on antigenic cross reactivity and sequence similarity HRV8 and HRV95 were combined into one single type [1]; however, since there are differences within the area equivalent to the receptor footprint they were considered separate types in the present communication. Like all other major group HRVs, K-types cannot infect via LDLR and/or LRP; they are neutralized by soluble ICAM-1 [7] and prevented from infecting HeLa cells by receptor

* Corresponding author. Address: Max F. Perutz Laboratories, University Departments at the Vienna Biocenter, Department of Medical Biochemistry, Medical University of Vienna, Dr. Bohr Gasse 9/3, A-1030 Vienna, Austria. Fax: +43 1 4277 9616.
E-mail address: dieter.blaas@meduniwien.ac.at (D. Blaas).

0014-5793/$36.00 © 2009 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2009.07.015

blockage with a monoclonal antibody against ICAM-1 [8]. Phylogenetically, minor group viruses form three subclusters and K-types are combined in two subclusters and two outliers, HRV24 and HRV58; when compared to the similarity between other HRVs the analysis does not suggest a higher phylogenetic relationship between the clusters [3]. Only recently the basis of receptor specificity emerged as a combination of charge complementarity and hydrophobic interactions [9]. Reasoning that the 12 minor group HRVs are recognized by the same receptor, presumably via interactions that differ for each type, we attempted to distinguish the two rhinovirus groups (and in particular minor group and K-type HRVs) by using a simple, largely automatable and unbiased bioinformatic approach; the 3D-structures of the contact sites between receptor and VP1 were modelled for all rhinoviruses based on the protein sequences [1] and the theoretical binding energies were calculated by three different approaches. The best performing method correctly classified K-type and minor group HRVs with all the latter exhibiting higher calculated affinities towards the receptor. This demonstrates the utility of energy calculations for the identification of binding partners by using 3D homology models.

2. Methods

2.1. Modelling VP1

VP1 sequences of the 101 HRVs [1] were downloaded from the UniProt knowledgebase. Note that HRV87 is identical with the acid-sensitive enterovirus EV68 [10] and therefore it was not considered further. The sequences aligned with ClustalW were truncated by removal of 70 residues from the N-terminus and as many from the C-terminus as to leave 180 residues. The resulting sequences were submitted to SwissModel [11] in 'first approach mode' with default parameters by using a PERL script for automation. Except from HRV7 and HRV69, sound models were obtained for all HRVs. For the latter two, visual inspection revealed that the program could not correctly build the loops not even when the sequences were resubmitted to SwissModel in 'optimized project mode' [12] by using HRV14 as template. Therefore, they were excluded from further analysis. Finally, 3D models of all 98 VP1 proteins including the BC, DE, and HI loops making up the receptor-binding epitope were obtained.

2.2. Modelling VP1–VP1′–V3

Since the footprint of receptor module V3 extends over two symmetry-related copies of VP1, such VP1–VP1* 'dimers' were assembled by superposition onto the experimental structure of VP1–VP1* of HRV2 by the 'magic fit' routine in Swiss-Pdb Viewer (SDBV4.0; Ref. [13]) by using a script. The VP1-dimers were energy minimized (100 cycles, steepest descendent minimization method) and final structures of the respective VP1–VP1*–V3 complexes obtained by combination with the coordinates of V3 taken from the HRV2-V23 X-ray structure [6]. Coordinates of these complexes were again energy-minimized as above. Note that the Ca^{2+} was not considered because no force filed parameters were available in *SPDBV*. Modelling of the HRV70 and HRV91 receptor complexes did not result in reasonable structures as their BC loops clashed with V3; this problem was not solved by energy minimization. As they are typical major group HRVs and not K-type viruses, we made no further effort to improve the models and excluded them from further analyses.

2.3. Energy calculations

Models of VP1–VP1*–V3 were submitted to the *Dcomplex* [14] web server (http://sparks.informatics.iupui.edu/song/complex.html). Data were entered and results retrieved automatically by using a PERL script. Models were also submitted to the *FastContact2.0* web server (http://structure.pitt.edu/servers/fastcontact/) [15,16] manually entering and retrieving the data. A local copy of the *FastContact2.0* program, kindly provided by Carlos Camacho, was employed as well. This latter software does not include the *CHARMm19* minimizer, which is being used in the web based version.

3. Results and discussion

As known from the 3D structure of the complex between V23, a two-module fragment of human VLDLR, and HRV2, the receptor interacts with VP1 only. We thus limited our model building efforts to module V3 of the receptor and the latter viral capsid protein. To reduce calculation time, the first 70 residues of all aligned VP1 proteins were removed and only the next 180 residues were considered. The deleted amino acids do not take part in the interaction and are even not involved in extensive contacts with the symmetry related VP1*. Templates selected by *SwissModel* running in 'automatic mode' are listed in Table 1. In accordance with the phylogenetic relationships [5] the program automatically selected the PDB coordinates of the B-type viruses HRV3 and HRV14 as templates for modelling of VP1 of the other B-types. For modelling the A-types, coordinates of HRV1A, HRV2, or HRV16 were automatically chosen as templates. Regarding the receptor groups there was no particular preference of the minor group types for any of the type-A templates, whereas for the K-types only HRV1A was used. As expected, for those HRVs whose 3D coordinates were in the database (HRV1A, 2, 3, 14, and 16) the corresponding data were selected for model building. To assess the reliability of the approach, VP1 from the latter viruses with available structures were also modelled automatically by excluding their own coordinates as templates. As seen in Table 2, all the models were within less than 0.65 Å root mean square deviation (RMSD) for the backbone (and less than 0.74 Å with the side chains included) from the experimental structure indicating good quality of the models.

3.1. FastContact performs better than Dcomplex in calculation of the binding energies

Having verified that our approach resulted in 3D models very well matching the known X-ray structures, we assumed that the other models were plausible and close to reality. Thus, we next

Table 1
Templates automatically selected by SwissModel for modelling VP1 of all HRVs. Blue, minor group; red, genus A major group; green, K-types (all genus A); orange, genus B major group; grey, HRV70 and HRV91 whose modelling led to strong crashes with V3. Striped, HRVs whose 3D X-ray structures are available. Note that in case of HRV48 and HRV72 the structure of HRV14 containing the antiviral capsid-binding hydrophobic antiviral compound WIN 52084 was automatically selected for modelling (accession number 1rud).

HRV1A 1r1a				HRV 2 1fpn	HRV 3 1rhi	HRV14 1k5m	HRV16 1aym
1A	95	39	65	2	70	5	25
1B	98	41	66	23	91	6	62
29	11	43	68	30	3	14	10
31	12	46	71	49	4	27	16
44	13	50	73	9	17	37	21
47	15	51	74	32	26	48*	45
8	19	53	75	67	35	52*	77
18	20	55	76		42	72*	81
24	22	57	78		79	84	90
40	28	59	80		83	86	96
54	33	60	82		92	93	100
56	34	61	88		99	97	HANKS
58	36	63	89				
85	38	64	94			*1rud	

Table 2
Quality of the 3D modelling. Models of HRVs whose 3D structure is known were automatically built by excluding themselves as template. Sequences were submitted to "template search" in SwissModel and the "second best template" (squared brackets) was chosen for modelling using "optimized project mode". RMS was derived from superposition of the models and the corresponding experimental structures by using "iterative magic fit" in SPDBV4.0. RMS, root mean square differences between coordinates of experimental and modelled structure.

Type	Used template	RMS Cα (Å)	RMS backbone + side chains (Å)
HRV1A	1aym [HRV16]	0.64	0.73
HRV2	1aym [HRV16]	0.57	0.60
HRV3	1k5 m [HRV14]	0.50	0.53
HRV14	1rhi [HRV3]	0.50	0.55
HRV16	1r1a [HRV1A]	0.61	0.70

asked whether the binding energies calculated for the complexes between receptor module V3 and the respective virus correlate with virus classification [8,17]. All models were either energy-minimized with the *SPDBV*-inbuilt *GROMOS96* or with *Jackal* [18] and subjected to energy calculation via *FastContact2.0* and *Dcomplex*. Note that *CHARMm19* is part of the web-based *FastContact* program; therefore, in order to prevent additional minimization with *CHARMm*, a local copy of *FastContact* that does not include the minimizer, was also used (in this case the structures were only minimized within *SPDBV*). The predicted affinities are summarized in Fig. 1. Overall, major and minor group viruses were quite well separated by *FastContact* regardless of the minimization method used (Fig. 1A–C). Nevertheless, the respective presence and absence of sequence motives in VP1 and VP3 within the ICAM-1 footprint [5] distinguishes these groups and can reliably used for classification of field isolates. However, when taking separation of minor group HRVs from K-types as a criterion, *FastContact* run on the webserver clearly gave the best results (Fig. 1A). Yet, as seen from the two delimiting lines, two K-type viruses (HRV18 and HRV98) and three minor group viruses (HRV2, HRV47, and HRV49) came very close to each other not allowing for their unambiguous classification. The separation of minor group HRVs and K-types was heavily compromised when energy minimization by *SPDBV* was omitted and carried out by the web-based *FastContact* alone (Fig. 1B). Prior minimization with *Jackal* did not improve this situation (Fig. 1C) and *Dcomplex* was even unable to grossly separate the minor and the major group, neither after minimization with *SPDBV* (Fig. 1D) nor with *Jackal*; by the same token, the local copy of *FastContact* did not separate the virus groups suggesting that additional minimization with *CHARMm* is important (data not shown).

3.2. Different contribution of amino acid residues to the binding energy

FastContact not only calculates the free energy from electrostatics, desolvation, and van der Waals contacts (the latter values were all below −500 kcal/mol making clashes unlikely; these values are not being used in the final energy calculation [15]), but also outputs the twenty residue pairs with lowest (attraction) and highest energy values (repulsion). In order to study which of these pairs might be most important in discrimination of minor group and K-type HRVs, the individual contributions to the total free interaction energy of the first ten were presented in the form of a heat map (data not shown). As expected, even in the K-type viruses the lysine in the HI loop emerged as the main player by interacting

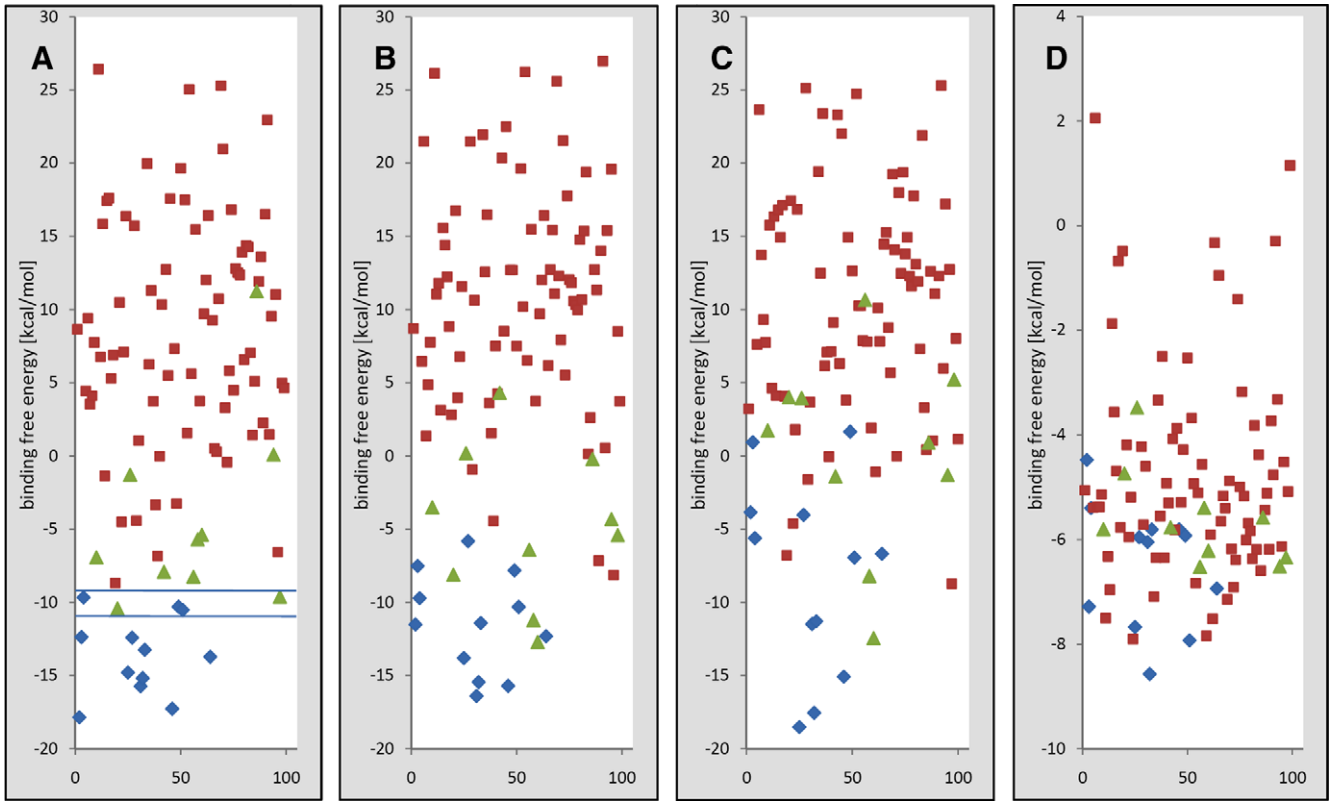

Fig. 1. Net interaction free energy (sum of the desolvation and the electrostatic energy) of the binding of V3 to 97 HRV types, listed on the horizontal axis, calculated with (A) *FastContact* webserver after energy minimization with *SPDBV3.7*, (B) *FastContact* webserver without energy minimization with *SPDBV3.7*, (C) *FastContact* webserver after energy minimization with *Jackal*, and (D) *Dcomplex* webserver. Major (red), minor (blue), and K-type viruses (green) are positioned according to their number. Note that in (A) all non-K-type major group viruses are well separated from the minor group. However, the K-types HRV18 and 98 and the minor group viruses HRV2, HRV47, and HRV49 (between the two lines) exhibit almost the same predicted binding affinities.

with glutamic and aspartic acid residues in the acidic cluster of the receptor module. In addition, various other residues were seen to differently contribute. Closer inspection drew our attention to the first cysteine in the receptor module because unexpectedly it was in the list of interacting residues. This cysteine forms a disulphide bridge with the 3rd cysteine and its sulfur atom is within 4 Å from the $-NH_2$ of lysine 228 of VP1 in HRV2 [6]. *FastContact* also sends back the coordinates corresponding to the energy-minimized structure used for calculation of the binding energies; this enabled us to visually inspect the data. It revealed that all disulphide bridges had been opened resulting in free –SH groups that had moved and presumably they were thus able to establish hydrogen bonds with suitable oxygen atoms nearby which were taken into account in the energy calculations. For the reasons above, this is not meaningful and can be definitely excluded. Therefore, we subtracted the energy values contributed by the two cysteines from the absolute sum of the interaction energies (see Fig. 1S for the heat maps with the corrected values). Energies of the receptor residues contributing to the interaction, except from the cysteines, are also depicted in the form of bar diagrams in Fig. 2S. Note that in many cases one given residue of the receptor interacts with more than one viral residue (compare to the matrix presentation in Fig. 1S). Although a clear cut difference in these patterns between the minor group and the K-types was not evident, it is obvious that involvement of the receptor residues is not equal given that the different virus types possess different residues at the equivalent positions (compare to Fig. 1S). For statistical analysis the mean of the free energies contributed by each receptor residue was computed for both virus groups and the standard error was used for estimation of significance by the *t*-test (Fig. 2). Most importantly, higher energy values of the K-type viruses for the interaction of the lysine with the acidic cluster were found, suggesting that different overall conformation sterically impedes optimal contacts of the respective residues. In Fig. 3, we replotted the affinities calculated for the minor group and the K-type viruses in the same way as in Fig. 1 but also corrected by subtraction of the values of the cysteines. This led to some change in the positions of the virus types in the diagram and moved the K-type HRV8 closer to the minor group HRVs. However, all minor group viruses were now separated from the K-types and thus correctly classified. This demonstrates that an approximate prediction of the relative affinities from 3D models is possible although the absolute values for HRV2 are far from those measured by fluorescence correlation spectroscopy [19]. As seen from the significant difference in the free energies contributed by the acidic cluster at the C-terminus of the receptor modules, electrostatics appears to be the major factor in discrimination. This is most obvious for HRV85, where the strong repulsive force between E25 in the receptor and D89 in the virus (see Figs. 1S and 2S) most probably prevents binding. The predominant role of charge complementarity in mutual recognition has already been proposed earlier [9,20]; it might be one of the reasons for the correct classification even when disregarding the van der Waals forces (see above). We are aware that other energy calculation programs are available; our choice was driven by the result of a web search in which the two used here showed up as major hits. Although we cannot exclude that even better results might be obtained by using other software, model building and energy calculation as used here astonishingly well identified specific characteristics of minor group and K-type HRVs that are suggested from visual inspection but could so far only be detected by experimentation.

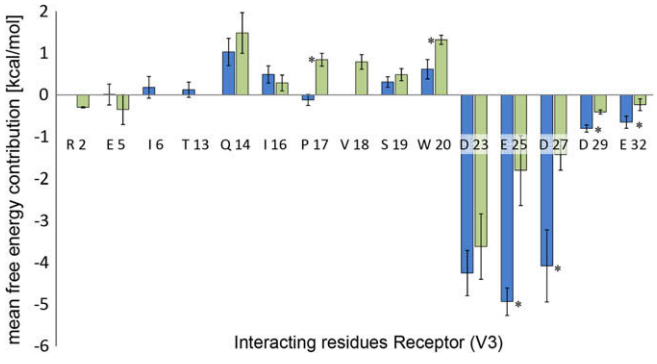

Fig. 2. Comparison of the mean interaction energies of amino acid residue of V3 interacting with residues of the minor group (blue) and K-type viruses (green) as calculated with *Fastcontact* (webserver version). The mean energy values for each residue (as shown individually in Fig. 2S) and the standard error of the mean (error bars) are given for minor group and K-type viruses as indicated. Significant differences (∗) for a 95% confidence interval were estimated by the *t*-test. Note that only those receptor residues found to contribute to the interaction in more than three virus types of the same group (i.e. being present in the respective list of the first 10 highest and lowest energy scores) are shown.

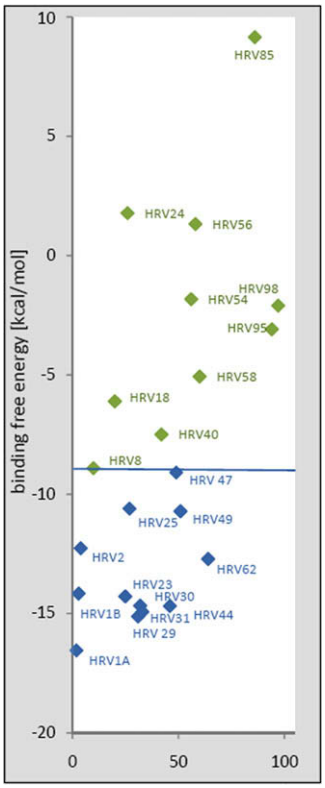

Fig. 3. Net interaction energies of the binding of V3 to minor group and K-type HRVs calculated with the webserver version of *Fastcontact* as in Fig. 1A but corrected for the erroneous contribution of the cysteines. Minor (blue), and K-type viruses (green) are positioned according to their number. Note that all minor group and K-type viruses are well separated as illustrated by the separation line.

Acknowledgements

Funded by the Austrian Science Foundation (FWF Grant # P18693-B09) and by the BIN-II and BIN-III projects of the Austrian GEN-AU programme. S.S. would like to thank the WWTF for generous funding.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2009.07.015.

References

[1] Ledford, R.M., Patel, N.R., Demenczuk, T.M., Watanyar, A., Herbertz, T., Collett, M.S. and Pevear, D.C. (2004) VP1 sequencing of all human rhinovirus serotypes: insights into genus phylogeny and susceptibility to antiviral capsid-binding compounds. J. Virol. 78, 3663–3674.

[2] McErlean, P., Shackelton, L.A., Andrews, E., Webster, D.R., Lambert, S.B., Nissen, M.D., Sloots, T.P. and Mackay, I.M. (2008) Distinguishing molecular features and clinical characteristics of a putative new rhinovirus species, human rhinovirus C (HRV C). PLoS ONE 3, e1847.

[3] Palmenberg, A.C., Spiro, D., Kuzmickas, R., Wang, S., Djikeng, A., Rathe, J.A., Fraser-Liggett, C.M. and Liggett, S.B. (2009) Sequencing and analyses of all known human rhinovirus genomes reveal structure and evolution. Science 324, 55–59.

[4] Olson, N.H., Kolatkar, P.R., Oliveira, M.A., Cheng, R.H., Greve, J.M., McClelland, A., Baker, T.S. and Rossmann, M.G. (1993) Structure of a human rhinovirus complexed with its receptor molecule. Proc. Natl. Acad. Sci. USA 90, 507–511.

[5] Laine, P., Blomqvist, S., Savolainen, C., Andries, K. and Hovi, T. (2006) Alignment of capsid protein VP1 sequences of all human rhinovirus prototype strains: conserved motifs and functional domains. J. Gen. Virol. 87, 129–138.

[6] Verdaguer, N., Fita, I., Reithmayer, M., Moser, R. and Blaas, D. (2004) X-ray structure of a minor group human rhinovirus bound to a fragment of its cellular receptor protein. Nat. Struct. Mol. Biol. 11, 429–434.

[7] Crump, C.E., Arruda, E. and Hayden, F.G. (1993) In vitro inhibitory activity of soluble ICAM-1 for the numbered serotypes of human rhinovirus. Antivir. Chem. Chemother. 4, 323–327.

[8] Uncapher, C.R., Dewitt, C.M. and Colonno, R.J. (1991) The major and minor group receptor families contain all but one human rhinovirus serotype. Virology 180, 814–817.

[9] Querol-Audi, J., Konecsni, T., Pous, J., Carugo, O., Fita, I., Verdaguer, N. and Blaas, D. (2009) Minor group human rhinovirus–receptor interactions: geometry of multimodular attachment and basis of recognition. FEBS Lett. 583, 235–240.

[10] Blomqvist, S., Savolainen, C., Raman, L., Roivainen, M. and Hovi, T. (2002) Human rhinovirus 87 and enterovirus 68 represent a unique serotype with rhinovirus and enterovirus features. J. Clin. Microbiol. 40, 4218–4223.

[11] Schwede, T., Kopp, J., Guex, N. and Peitsch, M.C. (2003) SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acid Res. 31, 3381–3385.

[12] Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18, 2714–2723.

[13] Guex, N., Diemand, A. and Peitsch, M.C. (1999) Protein modelling for all. Trends Biochem. Sci. 24, 364–367.

[14] Liu, S., Zhang, C., Zhou, H. and Zhou, Y. (2004) A physical reference state unifies the structure-derived potential of mean force for protein folding and binding. Proteins 56, 93–101.

[15] Champ, P.C. and Camacho, C.J. (2007) FastContact: a free energy scoring tool for protein-protein complex structures. Nucleic Acid Res. 35, W556–560.

[16] Camacho, C.J. and Zhang, C. (2005) FastContact: rapid estimate of contact and binding free energies. Bioinformatics 21, 2534–2536.

[17] Vlasak, M., Roivainen, M., Reithmayer, M., Goesler, I., Laine, P., Snyers, L., Hovi, T. and Blaas, D. (2005) The minor receptor group of human rhinovirus (HRV) includes HRV23 and HRV25, but the presence of a lysine in the VP1 HI loop is not sufficient for receptor binding. J. Virol. 79, 7389–7395.

[18] Petrey, D., Xiang, Z., Tang, C.L., Xie, L., Gimpelev, M., Mitros, T., Soto, C.S., Goldsmith-Fischman, S., Kernytsky, A., Schlessinger, A., Koh, I.Y., Alexov, E. and Honig, B. (2003) Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling. Proteins 53 (Suppl. 6), 430–435.

[19] Wruss, J., Runzler, D., Steiger, C., Chiba, P., Kohler, G. and Blaas, D. (2007) Attachment of VLDL receptors to an icosahedral virus along the 5-fold symmetry axis: multiple binding modes evidenced by fluorescence correlation spectroscopy. Biochemistry 46, 6331–6339.

[20] Vlasak, M., Blomqvist, S., Hovi, T., Hewat, E. and Blaas, D. (2003) Sequence and structure of human rhinoviruses reveal the basis of receptor discrimination. J. Virol. 77, 6923–6930.